

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/533066  
Source: PCT  
Date Processed by STIC: 5/10/5

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 05/10/2005

PATENT APPLICATION: US/10/533,066

TIME: 13:47:34

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05102005\J533066.raw

```

3 <110> APPLICANT: Takeda Pharmaceutical Compnay Limited
5 <120> TITLE OF INVENTION: Use of SGLT homolog
7 <130> FILE REFERENCE: G05-0003
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/533,066
C--> 9 <141> CURRENT FILING DATE: 2005-04-28
9 <150> PRIOR APPLICATION NUMBER: JP 2002-314041
10 <151> PRIOR FILING DATE: 2002-10-29
12 <150> PRIOR APPLICATION NUMBER: JP 2003-156306
W--> 13 <151> PRIOR FILING DATE: 2003-6-2
15 <160> NUMBER OF SEQ ID NOS: 57
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 674
19 <212> TYPE: PRT
20 <213> ORGANISM: Homo Sapiens
22 <400> SEQUENCE: 1
23 Met Gly Pro Gly Ala Ser Gly Asp Gly Val Arg Thr Glu Thr Ala Pro
24           5              10              15
25 His Ile Ala Leu Asp Ser Arg Val Gly Leu His Ala Tyr Asp Ile Ser
26           20              25              30
27 Val Val Val Ile Tyr Phe Val Phe Val Ile Ala Val Gly Ile Trp Ser
28           35              40              45
29 Ser Ile Arg Ala Ser Arg Gly Thr Ile Gly Gly Tyr Phe Leu Ala Gly
30           50              55              60
31 Arg Ser Met Ser Trp Trp Pro Ile Gly Ala Ser Leu Met Ser Ser Asn
32           65              70              75              80
33 Val Gly Ser Gly Leu Phe Ile Gly Leu Ala Gly Thr Gly Ala Ala Gly
34           85              90              95
35 Gly Leu Ala Val Gly Gly Phe Glu Trp Asn Ala Thr Trp Leu Leu Leu
36           100             105             110
37 Ala Leu Gly Trp Val Phe Val Pro Val Tyr Ile Ala Ala Gly Val Val
38           115             120             125
39 Thr Met Pro Gln Tyr Leu Lys Lys Arg Phe Gly Gly Gln Arg Ile Gln
40           130             135             140
41 Val Tyr Met Ser Val Leu Ser Leu Ile Leu Tyr Ile Phe Thr Lys Ile
42           145             150             155             160
43 Ser Thr Asp Ile Phe Ser Gly Ala Leu Phe Ile Gln Met Ala Leu Gly
44           165             170             175
45 Trp Asn Leu Tyr Leu Ser Thr Gly Ile Leu Leu Val Val Thr Ala Val
46           180             185             190
47 Tyr Thr Ile Ala Gly Gly Leu Met Ala Val Ile Tyr Thr Asp Ala Leu
48           195             200             205
49 Gln Thr Val Ile Met Val Gly Gly Ala Leu Val Leu Met Phe Leu Gly
50           210             215             220

```

## RAW SEQUENCE LISTING

DATE: 05/10/2005

PATENT APPLICATION: US/10/533,066

TIME: 13:47:34

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05102005\J533066.raw

```

51 Phe Gln Asp Val Gly Trp Tyr Pro Gly Leu Glu Gln Arg Tyr Arg Gln
52 225                230                235                240
53 Ala Ile Pro Asn Val Thr Val Pro Asn Thr Thr Cys His Leu Pro Arg
54                245                250                255
55 Pro Asp Ala Phe His Met Leu Arg Asp Pro Val Ser Gly Asp Ile Pro
56                260                265                270
57 Trp Pro Gly Leu Ile Phe Gly Leu Thr Val Leu Ala Thr Trp Cys Trp
58                275                280                285
59 Cys Thr Asp Gln Val Ile Val Gln Arg Ser Leu Ser Ala Lys Ser Leu
60                290                295                300
61 Ser His Ala Lys Gly Gly Ser Val Leu Gly Gly Tyr Leu Lys Ile Leu
62 305                310                315                320
63 Pro Met Phe Phe Ile Val Met Pro Gly Met Ile Ser Arg Ala Leu Phe
64                325                330                335
65 Pro Asp Glu Val Gly Cys Val Asp Pro Asp Val Cys Gln Arg Ile Cys
66                340                345                350
67 Gly Ala Arg Val Gly Cys Ser Asn Ile Ala Tyr Pro Lys Leu Val Met
68                355                360                365
69 Ala Leu Met Pro Val Gly Leu Arg Gly Leu Met Ile Ala Val Ile Met
70                370                375                380
71 Ala Ala Leu Met Ser Ser Leu Thr Ser Ile Phe Asn Ser Ser Ser Thr
72 385                390                395                400
73 Leu Phe Thr Ile Asp Val Trp Gln Arg Phe Arg Arg Lys Ser Thr Glu
74                405                410                415
75 Gln Glu Leu Met Val Val Gly Arg Val Phe Val Val Phe Leu Val Val
76                420                425                430
77 Ile Ser Ile Leu Trp Ile Pro Ile Ile Gln Ser Ser Asn Ser Gly Gln
78                435                440                445
79 Leu Phe Asp Tyr Ile Gln Ala Val Thr Ser Tyr Leu Ala Pro Pro Ile
80                450                455                460
81 Thr Ala Leu Phe Leu Leu Ala Ile Phe Cys Lys Arg Val Thr Glu Pro
82 465                470                475                480
83 Gly Ala Phe Trp Gly Leu Val Phe Gly Leu Gly Val Gly Leu Leu Arg
84                485                490                495
85 Met Ile Leu Glu Phe Ser Tyr Pro Ala Pro Ala Cys Gly Glu Val Asp
86                500                505                510
87 Arg Arg Pro Ala Val Leu Lys Asp Phe His Tyr Leu Tyr Phe Ala Ile
88                515                520                525
89 Leu Leu Cys Gly Leu Thr Ala Ile Val Ile Val Ile Val Ser Leu Cys
90                530                535                540
91 Thr Thr Pro Ile Pro Glu Glu Gln Leu Thr Arg Leu Thr Trp Trp Thr
92 545                550                555                560
93 Arg Asn Cys Pro Leu Ser Glu Leu Glu Lys Glu Ala His Glu Ser Thr
94                565                570                575
95 Pro Glu Ile Ser Glu Arg Pro Ala Gly Glu Cys Pro Ala Gly Gly Gly
96                580                585                590
97 Ala Ala Glu Asn Ser Ser Leu Gly Gln Glu Gln Pro Glu Ala Pro Ser
98                595                600                605
99 Arg Ser Trp Gly Lys Leu Leu Trp Ser Trp Phe Cys Gly Leu Ser Gly

```

## RAW SEQUENCE LISTING

DATE: 05/10/2005

PATENT APPLICATION: US/10/533,066

TIME: 13:47:34

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05102005\J533066.raw

```

100      610      615      620
101 Thr Pro Glu Gln Ala Leu Ser Pro Ala Glu Lys Ala Ala Leu Glu Gln
102 625      630      635      640
103 Lys Leu Thr Ser Ile Glu Glu Glu Pro Leu Trp Arg His Val Cys Asn
104      645      650      655
105 Ile Asn Ala Val Leu Leu Leu Ala Ile Asn Ile Phe Leu Trp Gly Tyr
106      660      665      670
107 Phe Ala
108      674
110 <210> SEQ ID NO: 2
111 <211> LENGTH: 2022
112 <212> TYPE: DNA
113 <213> ORGANISM: Homo sapiens
115 <400> SEQUENCE: 2
116 atggggcctg gagcttcagg ggacgggggtc aggactgaga cagctccaca catagcactg      60
117 gactccagag ttggtctgca cgctacgac atcagcgtgg tggatcatcta ctttgtcttc      120
118 gtcattgctg tggggatctg gtcgtccatc cgtgcaagtc gagggaccat tggcggctat      180
119 ttcctggccg ggagggtccat gagctgggtgg ccaattggag catctctgat gtccagcaat      240
120 gtgggcagtg gcttggtcat cggcctgggt gggacagggg ctgccggagg ccttgccgta      300
121 ggtggcttcg agtggaaagc aacctgggtg ctctggccc ttggtgggt cttcgtccct      360
122 gtgtacatcg cagcaggtgt ggtcacaatg ccgcagtatc tgaagaagcg atttgggggc      420
123 cagaggatcc aggtgtacat gtctgtcctg tctctcatcc tctacatctt caccaagatc      480
124 tcgactgaca tcttctctgg agccctcttc atccagatgg cattgggctg gaacctgtac      540
125 ctctccacag ggatcctgct ggtgggtgact gccgtctaca ccattgcagg tggcctcatg      600
126 gccgtgatct acacagatgc tctgcagacg gtgatcatgg tagggggagc cctggtcctc      660
127 atgtttcttg gctttcagga cgtggggtgg taccagggcc tggagcagcg gtacaggcag      720
128 gccatcccta atgtcacagt ccccaacacc acctgtcacc tcccacggcc cgatgctttc      780
129 cacatgcttc gggaccctgt gageggggac atcccttggc caggtctcat tttcgggctc      840
130 acagtgtctg ccacctgggtg ttggtgcaca gaccaggtca ttgtgcagcg gtctctctcg      900
131 gccaagagtc tgtctcatgc caaggagggc tccgtgctgg ggggctacct gaagatcctc      960
132 cccatgttct tcactgtcat gcctggcatg atcagccggg cctgttccc agacgaggtg      1020
133 ggtgcgtgg accctgatgt ctgccaaga atctgtgggg ccgagtgagg atgttccaac      1080
134 attgacctacc ctaagttggt catggccctc atgctgtttg gtctgcgggg gctgatgatt      1140
135 gccgtgatca tggccgctct catgagctca ctacacctca tcttcaacag cagcagcacc      1200
136 ctgttcacca ttgatgtgtg gcagcgtctc cgcaggaagt caacagagca ggagctgatg      1260
137 gtggtgggca gagtgtttgt ggtgttctcg gttgtcatca gcacctctcg gatccccatc      1320
138 atccaaagct ccaacagtgg gcagctcttc gactacatcc aggtgtcac cagttacctg      1380
139 gccccaccca tcaccgtctc ctctctgctg gccatcttct gcaagagggt cacagagccc      1440
140 ggagctttct ggggcctcgt gtttggcctg ggagtggggc ttctgcgtat gatcctggag      1500
141 ttctcatacc cagcgcacgc ctgtggggag gtggaccgga ggccagcagt gctgaaggac      1560
142 ttccactacc tgtactttgc aatcctctc tgcgggctca ctgccatcgt cattgtcatt      1620
143 gtcagcctct gtacaactcc catccctgag gaacagctca cagcctcac atggtggact      1680
144 cggaactgcc ccctctctga gctggagaag gagggccacg agagcacacc ggagatatcc      1740
145 gagaggccag ccggggagtg ccctgcagga ggtggagcgg cagagaactc gagcctgggc      1800
146 caggagcagc ctgaagcccc aagcaggtcc tggggaaagt tgctctggag ctggttctgt      1860
147 gggctctctg gaacaccgga gcaggccctg agcccagcag agaaggctgc gctagaacag      1920
148 aagctgacaa gcattgagga ggagccactc tggagacatg tctgcaacat caatgctgtc      1980
149 cttttgctgg ccataacat ctctctctgg ggctattttg cg      2022
151 <210> SEQ ID NO: 3

```

## RAW SEQUENCE LISTING

DATE: 05/10/2005

PATENT APPLICATION: US/10/533,066

TIME: 13:47:34

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05102005\J533066.raw

```

152 <211> LENGTH: 678
153 <212> TYPE: PRT
154 <213> ORGANISM: Mus musculus
156 <400> SEQUENCE: 3
157 Met Glu Pro Gly Val Ser Arg Asn Gly Val Arg Thr Glu Thr Thr Thr
158           5           10           15
159 Asn Pro Ser Leu Gly Leu His Thr Tyr Asp Ile Val Val Val Val Ile
160           20           25           30
161 Tyr Phe Val Phe Val Leu Ala Val Gly Ile Trp Ser Ser Ile Arg Ala
162           35           40           45
163 Ser Arg Gly Thr Val Gly Gly Tyr Phe Leu Ala Gly Arg Ser Met Thr
164           50           55           60
165 Trp Trp Pro Ile Gly Ala Ser Leu Met Ser Ser Asn Val Gly Ser Gly
166           65           70           75           80
167 Leu Phe Ile Gly Leu Ala Gly Thr Gly Ala Ala Gly Gly Leu Ala Val
168           85           90           95
169 Gly Gly Phe Glu Trp Asn Ala Thr Phe Leu Leu Leu Ala Leu Gly Trp
170           100          105          110
171 Ile Phe Val Pro Val Tyr Ile Ala Ala Gly Val Val Thr Met Pro Gln
172           115          120          125
173 Tyr Leu Lys Lys Arg Phe Gly Gly Gln Arg Ile Gln Val Tyr Met Ser
174           130          135          140
175 Val Leu Ser Leu Ile Leu Tyr Ile Phe Thr Lys Ile Ser Thr Asp Ile
176           145          150          155          160
177 Phe Ser Gly Ala Leu Phe Ile Gln Met Ala Leu Gly Trp Asn Leu Tyr
178           165          170          175
179 Leu Ser Thr Val Ile Leu Leu Val Val Thr Ala Val Tyr Thr Ile Ala
180           180          185          190
181 Gly Gly Leu Thr Ala Val Ile Tyr Thr Asp Ala Leu Gln Thr Val Ile
182           195          200          205
183 Met Val Gly Gly Ala Leu Val Leu Met Phe Leu Gly Phe Gln Glu Val
184           210          215          220
185 Gly Trp Tyr Pro Gly Leu Gln Gln Leu Tyr Arg Gln Ala Ile Pro Asn
186           225          230          235          240
187 Thr Thr Val Pro Asn Thr Thr Cys His Leu Pro Arg Pro Asp Ala Phe
188           245          250          255
189 His Met Leu Arg Asp Pro Val Asn Gly Asp Ile Pro Trp Pro Gly Leu
190           260          265          270
191 Ile Phe Gly Leu Thr Val Leu Ala Thr Trp Cys Trp Cys Thr Asp Gln
192           275          280          285
193 Val Ile Val Gln Arg Ser Leu Ala Ala Lys Asn Leu Ser His Ala Lys
194           290          295          300
195 Gly Gly Ser Val Leu Gly Gly Tyr Leu Lys Ile Leu Pro Met Phe Phe
196           305          310          315          320
197 Ile Val Met Pro Gly Met Ile Ser Arg Ala Leu Tyr Pro Asp Glu Val
198           325          330          335
199 Ala Cys Val Asp Pro Asp Ile Cys Gln Arg Val Cys Gly Ala Arg Val
200           340          345          350
201 Gly Cys Ser Asn Ile Ala Tyr Pro Lys Leu Val Met Ala Leu Met Pro

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/533,066

DATE: 05/10/2005

TIME: 13:47:34

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05102005\J533066.raw

```

202          355          360          365
203 Val Gly Leu Arg Gly Leu Met Ile Ala Val Ile Met Ala Ala Leu Met
204          370          375          380
205 Ser Ser Leu Thr Ser Ile Phe Asn Ser Ser Ser Thr Leu Phe Ala Ile
206 385          390          395          400
207 Asp Val Trp Gln Arg Phe Arg Arg Gln Ala Ser Glu Gln Glu Leu Met
208          405          410          415
209 Val Val Gly Arg Leu Phe Val Val Phe Leu Val Val Ile Ser Ile Leu
210          420          425          430
211 Trp Ile Pro Ile Ile Gln Ser Ser Asn Ser Gly Gln Leu Phe Asp Tyr
212          435          440          445
213 Ile Gln Ser Ile Thr Ser Tyr Leu Ala Pro Pro Ile Thr Ala Leu Phe
214          450          455          460
215 Leu Leu Ala Ile Phe Cys Lys Arg Val Asn Glu Pro Gly Ala Phe Trp
216 465          470          475          480
217 Gly Leu Met Phe Gly Leu Val Val Gly Ile Leu Arg Met Ile Leu Glu
218          485          490          495
219 Phe Ser Tyr Ser Ala Pro Ala Cys Gly Glu Met Asp Arg Arg Pro Ala
220          500          505          510
221 Val Leu Lys Asp Phe His Tyr Leu Tyr Phe Ala Leu Leu Leu Cys Gly
222          515          520          525
223 Leu Thr Ala Ile Ile Ile Val Val Ile Ser Phe Phe Thr Glu Pro Ile
224          530          535          540
225 Pro Asp Asp Lys Leu Ala Arg Leu Thr Trp Trp Thr Arg Asn Cys Ala
226 545          550          555          560
227 Val Ser Asp Leu Gln Lys Lys Thr Ser Val Ser Val Asn Asn Thr Glu
228          565          570          575
229 Asp Asp Asn Ser Pro Gly Leu Ala Gly Arg Pro Val Val Glu Gly Pro
230          580          585          590
231 Ala Gly Asp Glu Glu Glu Ala Asn Thr Thr Gln Gly Pro Glu Gln Pro
232          595          600          605
233 Gly Ala Leu His Arg Ser Trp Gly Lys Trp Leu Trp Asn Trp Phe Cys
234          610          615          620
235 Gly Leu Ser Gly Ala Pro Gln Gln Ala Leu Ser Pro Ala Glu Lys Ala
236 625          630          635          640
237 Val Leu Glu Gln Lys Leu Thr Ser Ile Glu Glu Glu Pro Leu Trp Arg
238          645          650          655
239 Arg Val Cys Asn Ile Asn Ala Ile Ile Leu Leu Ala Ile Asn Ile Phe
240          660          665          670
241 Leu Trp Gly Tyr Phe Ala
242          675          678
244 <210> SEQ ID NO: 4
245 <211> LENGTH: 2034
246 <212> TYPE: DNA
247 <213> ORGANISM: Mus musculus
249 <400> SEQUENCE: 4
250 atggaaccag gagtgtcaag gaatggagtc agaactgaga caacaacgaa cccaagcctg      60
251 gggctacata cctatgacat cgtgggtggtg gtcattctatt ttgtctttgt tcttgctgtg      120
252 ggaatttggt catccatccg tgcaagtcga gggaccgttg gtggctatit cctggctggg      180

```

**VERIFICATION SUMMARY**

DATE: 05/10/2005

PATENT APPLICATION: US/10/533,066

TIME: 13:47:35

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF4\05102005\J533066.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No  
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:13 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD  
L:460 M:283 W: Missing Blank Line separator, <400> field identifier  
L:470 M:283 W: Missing Blank Line separator, <400> field identifier  
L:524 M:283 W: Missing Blank Line separator, <400> field identifier